



DØnote 4399-CONF

## A Search for $Wb\bar{b}$ and $WH$ Production at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration

URL: <http://www-d0.fnal.gov>

(Dated: March 29, 2004)

A study of  $Wb\bar{b}$  production at  $\sqrt{s} = 1.96$  TeV in the  $e + \cancel{E}_T$  final state is presented for 174  $\text{pb}^{-1}$  of integrated luminosity. A good description of the events with 2  $b$ -tagged jets produced in association with a  $W$  candidate provides a 95 % C.L. upper limit on the  $Wb\bar{b}$  production of 20.3 pb. Although this limit is consistent with expectations from the standard model, the two events remaining in the exclusive subsample of  $e + \cancel{E}_T + 2$  tagged  $b$  jets appear to be excellent candidates for  $Wb\bar{b}$  production. With no events observed in a dijet mass window of 85–135 GeV, we set an upper limit on  $\sigma(p\bar{p} \rightarrow WH) \times B(H \rightarrow b\bar{b})$  of 12.4 pb at 95% C.L. for a Higgs boson mass of 115 GeV.

*Preliminary Results for Winter 2004 Conferences*

## I. INTRODUCTION

The  $Wb\bar{b}$  channel plays a special role at the Tevatron since it is an important background for Higgs and single-top production. This channel also provides a clean benchmark of the capabilities of the detector concerning jets, missing  $E_T$  ( $\cancel{E}_T$ ), lepton identification/calibration and  $b$ -tagging.

Our measurement is based on the “ $W + \geq 2$  jets” channel producing the  $e + \cancel{E}_T$  final state, to which different types of  $b$ -tagging can be applied. The properties of  $W +$  jets events are presented and compared to the simulated expectation. A study of the “ $W + \geq 2$  jets” channel producing the  $e + \cancel{E}_T$  final state, has been presented by DØ at the Lepton-Photon 2003 conference [1], in which a 95% C.L. upper limit on the  $Wb\bar{b}$  cross section of 33.4 pb was set based on 117.1 pb<sup>-1</sup> of data. In this note we extend this study by improving the previous cross section limit, selecting likely candidates of  $Wb\bar{b}$  production, and setting a cross-section limit on  $WH$  production.

## II. DATA SAMPLE AND EVENT SELECTION

The analysis relies on the following components of the DØ detector: a magnetic central-tracking system, which consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet [2]; a liquid-argon/uranium calorimeter made of a central section (CC) covering  $|\eta|$  up to  $\approx 1.1$ , and two end calorimeters (EC) extending coverage to  $|\eta| < 4.0$ , all housed in separate cryostats [3], with scintillators between the CC and EC cryostats providing sampling of developing showers at  $1.1 < |\eta| < 1.4$ ; a muon system which resides beyond the calorimetry, and consists of a layer of tracking detectors and scintillation trigger counters before 1.8 T toroids, followed by two more similar layers after the toroids; the tracking at  $|\eta| < 1$  relies on 10 cm wide drift tubes [3], while 1 cm mini-drift tubes are used at  $1 < |\eta| < 2$ .

The luminosity is measured using plastic scintillator arrays located in front of the EC cryostats, covering  $2.7 < |\eta| < 4.4$ . The trigger and data acquisition systems are designed to accommodate the large luminosity of Run II. We reject data periods in which the quality of the data of the tracking (CFT and SMT), the calorimeter or the muon system may be compromised. The resulting luminosity (174 pb<sup>-1</sup>) is about 10% lower than the total recorded luminosity during this period. The uncertainty on the measured luminosity is 6.5% [4].

The  $W +$  jets candidate events must pass one of the triggers requiring at least one EM object and are preselected by requiring at least one candidate electron (electromagnetic (EM) fraction  $> 0.9$ , isolation  $< 0.15$ , loose shower shape requirements, a loose matching to a track having  $p_T > 10$  GeV) with  $p_T^e > 12$  GeV and one jet satisfying standard DØ quality criteria with  $p_T^{jet} > 15$  GeV.

The calorimeter data are zero-suppressed (at  $2.5\sigma$ , where  $\sigma \equiv$  RMS of the cell noise), and in the event reconstruction we use a calorimeter reconstruction algorithm, so-called T42 [5, 6], which suppress all cells with negative energy and single isolated cells between  $2.5$  and  $4\sigma$  before the other calorimetric algorithms are applied. This has been shown to improve the calorimeter performance [7].

The event selection for this analysis requires at least one electron with  $p_T > 20$  GeV, missing transverse momentum  $\cancel{E}_T > 25$  GeV and at least one jet with  $p_T > 20$  GeV. Only events having a  $z$  vertex within  $\pm 60$  cm of the nominal interaction point are kept. The electron is required to be in the central region, i.e. to have  $|\eta_{detector}| < 1.1$ , and to be well contained, i.e. away from the module boundaries in  $\varphi$  [8].

The electrons are identified in 2 steps: a) the preselected electron candidates are first required to pass stricter identification criteria (electromagnetic (EM) fraction  $> 0.9$ , isolation  $< 0.10$ , stricter shower shape requirements, tighter track-cluster matching). These criteria define a “loose” electron. b) The loose electrons are then tested with a likelihood algorithm [9] developed on well controlled samples, and which takes as input 8 quantities sensitive to the EM nature of the particles. If they satisfy the likelihood requirement, they become final electrons for the analysis.

The inefficiencies induced by the ID and likelihood requirements are determined from a dielectron sample in which we select a pure set of  $Z$  events. The ID/reconstruction efficiency is found to be  $81\% \pm 3\%$ . The likelihood efficiency for electrons in the central region is determined to be  $91\% \pm 3\%$ . The invariant mass distribution for the  $Z \rightarrow ee$  events which are used to determine these efficiencies is shown in figure 1. It is compared to a detailed detector simulation of  $Z \rightarrow ee$  events, in which a rescaling by  $+0.75\%$  and an additional smearing of 3.5 % have been applied to the electron energy. These small ad hoc corrections are used in the analysis.

To estimate the number of multijets (QCD) events containing a jet passing the final electron identification criteria we determine from the data the probability  $p_{tight}^{loose}$  for a “loose” electron originating from a jet to pass the likelihood cut. This is done using a sample of events having 2 jets back to back in  $\varphi$  ( $|\pi - \Delta\varphi| < 0.2$ ), at low  $\cancel{E}_T (< 10$  GeV), and in which one of the jets has an EM fraction smaller than 0.7, is required to be in the central calorimeter and is far from the modules boundaries. The other jet is required to satisfy all the electron-id requirements except the likelihood cut. The probability  $p_{tight}^{loose}$  is obtained by dividing the number of events containing at least one electron candidate passing the likelihood cut by the total number of events of the sample. This probability is determined as a function of the  $p_T$

of the candidate electron. The QCD background is then estimated for every differential distribution: we use this  $p_T$  dependent probability in the so-called matrix method [24] that we apply to our final sample and to the loose sample (defined as the sample in which all the event selection criteria are applied but for the electron likelihood). The  $p_T$  distribution of the electron in the final  $W$ -Dijet sample is shown in figure 1, and compared to the expectation: at low  $p_T$  the contamination of QCD background as determined by the matrix method is visible. The shape and magnitude of the distribution is well reproduced by the ALPGEN simulation of the  $W$  + jets processes, after adding the QCD background and other standard model (SM) backgrounds detailed in the next section.

Since we want to select  $W$  decays we require large transverse missing energy:

$$\cancel{E}_T > 25\text{GeV}$$

$\cancel{E}_T$  is calculated from the calorimetric cells but for unclustered cells in the Coarse Hadronic [5], and is corrected for the presence of any muons. All energy corrections to electrons or to jets are propagated into  $\cancel{E}_T$ . The systematic uncertainties associated with the  $\cancel{E}_T$  selection criteria is obtained by varying the jet and electron energy scale within their systematics error,  $\pm 6\%$  and  $\pm 2\%$  respectively.

Small anisotropies of the calorimetric noise varying over time are taken into account: the missing  $E_T$  receives in average a contribution from the noise of about 1 GeV in the  $x$  and  $y$  component of  $\cancel{E}_T$ , and this contribution varies in amplitude and direction, such that integrated over a long period of time, this contribution is roughly Gaussian. To take into account in the simulation this effect, we add to  $\cancel{E}_{Tx}$  ( $\cancel{E}_{Ty}$ ) a Gaussian noise component having 0.0 GeV (-0.4 GeV) mean and 1.0 (1.2) GeV RMS. The resulting simulated distribution is compared to the data in figure 1c: the transverse missing energy distribution is well reproduced by the simulation when taking into account the QCD background which accumulates at low  $\cancel{E}_T$ .

The transverse mass of the  $W$  candidates in the  $W$ -Dijet sample is reconstructed from the electron and missing transverse energy. Its distribution is shown in figure 1d. A good agreement in shape and amplitude is observed between the data and the simulation.

### III. SIMULATED DATASETS

The following processes, which are used to compare with data, have been simulated with the PYTHIA [10] MC event generator version 6.202, making use of the CTEQ5L [11] leading-order parton distribution functions:

- inclusive production of:  $W \rightarrow e\nu$ ;  $Z \rightarrow ee$ ;  $W \rightarrow \tau\nu$ ;  $Z \rightarrow \tau\tau$ ;  $WZ$  production followed by  $Z \rightarrow b\bar{b}$ .
- $t\bar{t} \rightarrow e + \text{jets}$  production (lepton+jet and dilepton channels)
- $WH \rightarrow e\nu b\bar{b}$  production

The following processes are simulated using other generators:

- The single-top ( $s$ -channel ( $tb$ ) and  $tbq$ -channel) are generated using COMPHEP [12].
- The  $W + \geq 2$  jets events (“ $W$  + jets”), are generated with ALPGEN [13] (with PYTHIA radiation and showering) since ALPGEN has a better simulation of processes with large jet multiplicities. The generation is based on  $Wjj$  and  $Wjjj$  processes, including  $Wc\bar{c}$  and  $Wcj$ , but not  $Wb\bar{b}$  which is generated separately.
- The  $Wb\bar{b}$  events are generated with ALPGEN requiring 2 parton jets with  $p_T > 8$  GeV separated in  $\eta$ - $\varphi$  by  $\Delta R(\equiv \sqrt{\Delta\eta^2 + \Delta\varphi^2}) > 0.4$ .

The effect of PDF’s variation has been studied in Ref. [14] and a variation of the cross sections between 4 and 10% has been observed, depending on the process. We assume conservatively a systematic error on these cross sections of 15%. All the above events were processed through DØ detector simulation (DØgstar [15], based on GEANT), the electronics simulation (DØsim) and the reconstruction software (DØreco). The simulated events are then weighted by the trigger efficiency and by the ratio of the electron id and likelihood efficiencies for data and simulation ( $0.99 \pm 0.04$ ) These simulated backgrounds are absolutely normalized, i.e. according to cross section. In the sample “ $W + \geq 1$  jet”, the ratio between the total number of events in data and the number of simulated events (using PYTHIA), after efficiency correction, is equal to 1.08, which is consistent with 1, taking into account the systematic error of the DØ luminosity measurement (6.5%), and the typical error on the production cross sections for background ( $> 5 - 10\%$ ). This ratio is equal to 0.95 for the “ $W + \geq 2$  jets” when comparing to ALPGEN, and 1.48 when comparing to PYTHIA.

#### IV. KINEMATIC PROPERTIES OF $W$ -DIJET EVENTS

The jets used in this analysis are cone type proto-jets with a radius of  $R = 0.5$  satisfying the following cuts. These cuts are required to ensure that the jets are not due to rare spurious energy deposits.

- the energy fraction in the EM layers of the jet is required to be  $0.05 < EMF < 0.95$ .
- The energy fraction in the outermost layer (coarse hadronic) is required to be  $CHF < 0.4$ .
- Since the energy of the jet is expected to be shared between several calorimeter towers which have a size of  $0.1 \cdot 0.1$  in the  $(\eta, \varphi)$  plane, the number of towers containing 90% of the jet energy is required to be  $> 1$ .
- To avoid jets created from one single hot cell, we require the ratio of the energy in the hottest cell to the next-hottest to be  $< 10$ .
- The reconstructed  $p_T$  of the jet is compared to the equivalent quantity (L1.SET) built from the L1 trigger tower energies: the ratio  $L1.SET/p_T$  is required to be greater than 0.4 (0.2 in the  $0.8 < |\eta| < 1.5$  region).

If any of this cut is not satisfied, the energy of the proto-jet is simply counted in the unclustered energy. If a simulated event contains one or more reconstructed jets, the  $p_T$  dependent ratio of jet-id/reconstruction efficiency between data and simulation ( $0.95 \pm 0.07$  in average) is used (for every jet) to reweight the simulated events.

Since the expected number of  $Wb\bar{b}$  is small, we have first studied the kinematic properties of  $W + \geq 1$  jet events and found them well described by the simulation based on PYTHIA [16]. To test the quality of our simulation at higher jet rate we introduce the relative jet production rate defined as

$$R_{jets}^n \equiv \left( \frac{\text{number of jets in the data}}{\text{number of jets in the simulation}} \right)_{\text{in } W + \geq n \text{ jets events}}$$

For PYTHIA  $R_{jets}^1$  is found to be 1.08 and  $R_{jets}^2$  is equal to 1.48. The underproduction of events with jet multiplicity greater than 1 is understood since PYTHIA produces only the leading jet via matrix elements, the other jets being produced by parton shower in a “softer” way. For ALPGEN  $R_{jets}^2$  is equal to 0.95 so the predicted rate of  $W + \geq 2$  jets event is correct. In the following we use the ALPGEN sample, rescaled by a factor 0.95, for the  $W + \text{jets}$  expectation.

To check if the kinematic distributions of the jets are also well reproduced by the simulation, we study different jet characteristics: the distributions of the jet  $p_T$  and multiplicity are shown in figures 2a,b. The  $p_T$  distribution is reasonably well described by the simulation: the undershoot of the simulation in the lowest  $p_T$  bin is still in agreement with the data taking into account the 7% uncertainty on the jet id/reconstruction efficiency. The jet multiplicity is well described by the simulation up to the highest multiplicities. In each multiplicity bin, the composition of the simulation varies, providing confidence that the relative composition of the background is correctly described.

Kinematic properties of the jets are displayed in figures 2c,d which show the distance in  $R$  between the 2 leading jets, and the invariant mass of these 2 jets. The asymmetric shape of the  $\Delta R$  distribution is well described by the simulation. The dijet mass is well described for masses above 60 GeV.

In conclusion, the  $W$ -Dijet events are well described by the simulation which is mainly based on the ALPGEN + PYTHIA model for these  $W + \text{jets}$  processes. This gives confidence in the simulation when studying the properties of the  $Wb\bar{b}$  events, since this process is also simulated with the same ALPGEN + PYTHIA generator.

#### V. $b$ -TAGGING RESULTS

We concentrate in this section on searching for  $b$ -jet candidates in the  $W$ -Dijet sample. For tagging  $b$ -jets we use the JLIP algorithm [17]. JLIP is based on the estimation of the probability to observe the  $b$  lifetime. The jet lifetime probability is constructed using only the tracks associated to the jets which have a positive impact parameter in the transverse plane. The sign of the impact parameter is defined positive when the scalar product  $\vec{d}_{perigee} \cdot \vec{p}_T(jet)$  is positive.  $\vec{d}_{perigee}$  is defined in the transverse plane by the primary vertex and the impact point.

We cut on a jet lifetime probability smaller than 0.4% (tight cut), which is known to give a mistag rate (tagging of light quarks) of  $0.30 \pm 0.03\%$  [18]. For this mistag rate, the tagging efficiency for a “taggable” jet is expected to be  $39 \pm 2\%$ . A jet is “taggable” if at least 1 track jet is associated with it. The jet taggability is typically 80% in a single jet QCD sample [18].

The ratio between data and simulation for the taggability  $\times$  tagging efficiency ( $0.88 \times 0.72$ , as established from Ref. [18]) is used, for every  $b$ -tagged jet, to reweight the simulated events in which one or more jets originating from a  $b$  or a  $c$  are JLIP-tagged. In this way, the simulated distributions are corrected for the tagging efficiency and can

be directly compared to the data. The  $p_T$  and  $\eta$  dependence of the taggability and tagging efficiency, studied in [18] allow to derive a systematic error on the data/simulation ratio of these quantities of  $\pm 3\%$  and  $\pm 7\%$  respectively. No data/MC correction is applied for the tagging efficiency of light quarks, for which we assume a 20% error.

In the following, since we are studying  $b$ -tagged events, a  $\Delta R$  cut between the two leading jet is applied ( $\Delta R > 0.75$ ): this reduces the influence of  $c$  and  $b$  jets produced by gluon splitting and allows an unambiguous assignment in the simulation of the flavor of the jet. The simulated jet flavor is defined as the “highest” flavor found in a 0.3 cone centered around the direction of the reconstructed jet, the flavors being ranked from lowest to highest as  $u, d, s, c, b$ .

In figure 3a is shown the distributions for the  $p_T$  of the  $b$ -tagged jets for the  $W$ -dijet events having one JLIP  $b$ -tag jet(s): within errors, the data are well described by the simulation which includes the different standard model processes.

The multiplicity of  $b$ -tagged jets is shown in figure 3b. According to the simulation, many processes contribute in a significant way to the first bin, i.e. for the dijet events in which only one jet is  $b$ -tagged. Furthermore, the proportion of events with a mistagged jet is important in this first bin, as can be seen from the QCD and  $Wjj$  contribution. It is thus difficult to extract information on the processes containing genuine  $b\bar{b}$  in the final state, from single  $b$ -tagged events.

The kinematics of the single  $b$ -tagged events is well described by the simulation. In figure 3c and d are shown the  $\Delta R$  and dijet mass distributions in which one (and only one) jet has been tagged as a  $b$ -jet by JLIP. The data are compared to ALPGEN and to the other SM contributions.

We observe 100 events having one and only one  $b$ -tagged jet. The QCD background is estimated with the matrix method, using as loose sample the  $W$ -Dijet sample in which the electron is not required to pass the likelihood cut. It amounts to 18.5 events. The uncertainty on the  $p_{tight}^{loose}$  probability converts into an error of  $\pm 4.4$  events on this background.

For the simulation, the dominant JES error propagates into an error of  $22\% \times N_{SM}$ , where  $N_{SM}$  is the number of  $b$ -tagged events expected from standard model processes involving top quarks or weak interaction bosons events.

$N_{SM} = 69.5$  events when using ALPGEN for the simulation of  $W + n$  jets events. These numbers originates from:

- $44.8 \pm 14.9$  events due to  $W + \geq 2$  jets events
- $8.0 \pm 2.1$  events from  $Wb\bar{b}$  production
- $12.2 \pm 3.2$  events from  $t\bar{t}$  production
- $3.2 \pm 0.8$  events from single-top production
- $1.0 \pm 0.3$  events due to other  $W/Z$  decays ( $Z \rightarrow ee, Z \rightarrow \tau\tau, W \rightarrow \tau\nu$ )
- $0.3 \pm 0.1$  events from  $WZ$  production followed by  $Z \rightarrow b\bar{b}$

After QCD background subtraction ( $18.5 \pm 4.4$  events), we observe  $81.5 \pm 4.4$  single  $b$ -tag events to be compared to an expectation of  $69.5 \pm 21.4$ , in which the error is originating from:  $\pm 15.3$  (JES)  $\pm 4.5$  (Luminosity)  $\pm 10.4$  (cross section uncertainties)  $\pm 9.8$  (data/MC efficiencies). In conclusion, ALPGEN provides a good description of the jet properties of our sample and can be used to predict the expected rate of tagged events, within the systematic errors mentioned above.

## VI. DOUBLE $b$ -TAGGED EVENTS

When requesting only one  $b$ -tagged jet, the background due to  $W + 2$  light quark jets, top and QCD processes is still more than a factor 5 larger than the processes which have not yet been observed and which can be studied with the upgraded Tevatron:  $Wb\bar{b}$ , single-top or Higgs production. We thus study the events in which a second jet is  $b$ -tagged.

In figure 4a is shown the  $p_T$  distribution of the  $b$  tagged jets in events having at least 2  $b$ -tagged jets, compared to the simulated expectation. There are 17 data entries, corresponding to 7 events with 2  $b$ -tagged jets and one event with 3  $b$ -tagged jets. We expect that jets originating from  $Wb\bar{b}$  have in average a smaller  $p_T$  than those originating from  $t\bar{t}$  decays. The  $H_T$  distribution of these 8 events is shown figure 4b. 5 of them have a large  $H_T$  ( $> 180$  GeV), i.e in a region where the background is dominated by  $t\bar{t}$  events.

The distributions of  $\Delta R$  and invariant mass obtained from the two leading  $b$ -tagged jets are shown in figure 4c and d respectively. The 8 events observed are to be compared to an expected standard model background of  $8.3 \pm 2.2$  events:

The expected QCD background as obtained from the matrix method is  $0.7 \pm 0.2$  events, starting from the corresponding loose sample of 12 events. We expect  $1.4 \pm 0.3$  events from  $W + \geq 2$  jets as simulated by ALPGEN and  $0.2 \pm 0.06$  events from other  $W/Z$  decays.  $6.0 \pm 1.6$  events from processes with hard  $b$  production are also expected:

- $1.4 \pm 0.4$  events from  $Wb\bar{b}$ ,
- $4.0 \pm 1.0$  events from  $t\bar{t}$ ,
- $0.5 \pm 0.1$  events from single-top.
- $0.1 \pm 0.03$  events from  $WZ$  production followed by  $Z \rightarrow b\bar{b}$ .

We derive the propagation of the JES error in the 2  $b$ -tagged sample using the simulation: by varying the jet energy scale by  $\pm 1\sigma$ , the dijet rate varies, and we consider as JES error the relative variation of the doubly-tagged dijet rate normalized to the total dijet rate. This error amounts to  $\pm 14\%$ .

In summary, we observe 8 events for a total expectation of  $8.3 \pm 2.2$  events. All these results and distributions show that the simulation describes the data. To estimate our present sensitivity we derive a cross section limit for the  $Wb\bar{b}$  process.

## VII. $Wb\bar{b}$ CROSS SECTION LIMIT

In order to set a  $Wb\bar{b}$  or a  $WH$  cross section limit or to extract evidence for their production we first verify that the selected events are consistent with  $WX$  production. From the invariant transverse mass of the  $W$ -Dijet sample before tagging (figure 1d) we study the expected event distribution for  $Wjj$  events: after a sharp decrease for transverse masses between 80 and 40 GeV, we observe that below 30 GeV the rate of events increases, indicating that the background dominates over the  $W$  signal in this region. At high mass values, the expected transverse mass distribution shows that only a small contribution is expected above 120 GeV. Events appearing in this region come either from a strongly mismeasurement of  $\cancel{E}_T$ , or from processes having genuine  $\cancel{E}_T$ , beyond the  $\cancel{E}_T$  originating from the  $W$  that we are trying to isolate. To avoid these 2 types of events, we require the  $W$  candidates to have a transverse invariant mass between 25 and 125 GeV.

In figure 5a is shown the transverse invariant mass of the electron- $\cancel{E}_T$  system in events with 2  $b$ -tagged jets. Five events are well within a window of  $\pm 40$  GeV centered on the  $W$  mass, while 2 events are at low invariant transverse mass ( $< 20$  GeV) where some QCD background with mistagging is expected, and another one is at high mass (144 GeV). The dijet invariant masses for the selected events are shown in figure 5b, and compared to the expected number of  $6.9 \pm 1.8$  events, of which  $1.4 \pm 0.4$  events are expected from matrix element  $Wb\bar{b}$  production (ALPGEN), in which there are at least 2  $b$ -jets with energy greater than 8 GeV at the parton level, and  $0.7 \pm 0.3$   $Wb\bar{b}$  events as produced in the  $W + \geq 2$  jets ALPGEN + PYTHIA simulation. The other sources of background are  $t\bar{t}$  production ( $3.3 \pm 0.9$  events), single-top ( $0.5 \pm 0.1$  events), QCD background ( $0.4 \pm 0.1$  events), other  $W/Z$  decays ( $0.1 \pm 0.03$  events), and  $Wc\bar{c}$  production ( $0.4 \pm 0.1$  events) +  $Wq\bar{q}$  production ( $0.1 \pm 0.03$  events,  $q = u, d, s$ ) as generated by PYTHIA matched to ALPGEN.

The experimental systematic errors due to efficiencies (i.e. the error on the **ratio data/simulation of the efficiencies**) or to the propagation of other systematic errors (trigger, energy calibration, smearing), which affect the signal and standard model backgrounds (QCD background excepted) are the following:

- 3% error from the trigger efficiency; 3% error for the electron likelihood efficiency; 3% error for the electron-id/reconstruction efficiency
- 4% error due to the 2% systematic uncertainty on the electron energy; 1% error for the error due to the  $\pm 2\%$  uncertainty on the added 3.5% electron energy smearing.
- 7% for the jet-id and jet reconstruction efficiency.
- 15% due to the jet energy scale error, which includes the error on the  $\cancel{E}_T$  cut, due to the jet energy scale and to the  $\cancel{E}_T$  smearing.
- 4% for the jet taggability (3% per jet); 10% for the JLIP  $b$ -tagging efficiency (7% per jet)

giving an **overall experimental systematic error of 21%**. The uncertainty on the cross-sections used in the simulation is conservatively assumed to be 15%, hence the total systematic error on the simulated processes is 26%. The total number of expected “background” events, i.e. excluding  $Wb\bar{b}$  expected contribution, is  $4.8 \pm 1.2$ . The luminosity error is 6.5%.

TABLE I: Summary table for the  $e + \text{jets}$  final state: observed events in the data compared to the expected number of  $W + \geq 2$  jets before and after  $b$ -tagging, originating from QCD multijet background, top production ( $t\bar{t}$  and single-top) and “ $W + \text{jets}$ ” which includes  $W + \text{jets}$  in the final state, but also the small contributions of  $Z + \text{jets}$ , or  $WZ$  events which satisfy the selection criteria.  $Wb\bar{b}$  and  $Wc\bar{c}$  are also included in “ $W + \text{jets}$ ” and become strongly dominant once 2  $b$ -tagged jets are required, see text for details. The “triple-tag” is performed after application of the  $m_W^T$  cut ( $25 \text{ GeV} < M_W^T < 125 \text{ GeV}$ ), and requires that all jets of the event are tagged by the 3 DØ  $b$ -tagging algorithms. The “triple-tag” requirement is not used for the Higgs search.

	$W + \geq 2 \text{ jets}$	$W + \geq 2 \text{ jets}$ with 1 $b$ -tag	$W + \geq 2 \text{ jets}$ with 2 $b$ -tags	$W + \geq 2 \text{ jets}$ with 2 $b$ -tags and $m_W^T$ cut	$W + 2 \text{ jets}$ with 2 $b$ -tags and $m_W^T$ cut	$W + 2 \text{ jets}$ with 2 $b$ -tags and “triple-tag”
QCD Multijet	$497 \pm 119$	$18.5 \pm 4.4$	$0.7 \pm 0.2$	$0.4 \pm 0.1$	$0.0 \pm 0.1$	$0.0 \pm 0.1$
$t\bar{t} + \text{single-top}$	$51 \pm 13$	$15.4 \pm 4.0$	$4.5 \pm 1.2$	$3.8 \pm 1.0$	$0.6 \pm 0.2$	$0.2 \pm 0.1$
“ $W + \text{jets}$ ”	$2125 \pm 701$	$54.1 \pm 18.0$	$3.1 \pm 0.8$	$2.7 \pm 0.7$	$1.9 \pm 0.5$	$0.7 \pm 0.3$
Total expected	$2670 \pm 838$	$88.0 \pm 23.0$	$8.3 \pm 2.2$	$6.9 \pm 1.8$	$2.5 \pm 0.7$	$0.9 \pm 0.4$
Observed events	2567	100	8	5	2	2

The signal efficiency for the  $Wb\bar{b}$  signal is derived from the number of events (1.36) expected from the dedicated  $Wb\bar{b}$  ALPGEN simulation, and from the corresponding NLO cross section computed with MCFM [19] that we use for this  $Wb\bar{b}$  expectation:

$$\epsilon_{Wb\bar{b}} = 1.36 / (3.35 \text{ pb} \cdot 174 \text{ pb}^{-1}) = 0.00233 \pm 0.00049$$

The upper limit on the cross section is obtained using a Bayesian approach [20] that takes statistical and systematic uncertainties into account. **The 95% confidence level upper limit on the  $Wb\bar{b}$  cross section is 20.3 pb.** A summary of the number of expected events compared to the observed events is given in table I. The last 2 columns of the table are discussed in the next section.

### VIII. $Wb\bar{b}$ CANDIDATES AND $WH$ CROSS SECTION LIMIT

Since the dominant background to our signal is  $t\bar{t}$  production, we can straightforwardly improve our signal to background ratio by requiring exactly 2 jets in the final state, and no additional jet activity (low energy jets, or jets failing some identification criteria). After these conditions we are left with 2 events, as shown for the dijet invariant mass in figure 5c and d (linear and logarithmic scale). This is to be compared with  $(0.3 + 0.3) \pm 0.2$  events from  $t\bar{t} + \text{single-top}$  production,  $1.4 \pm 0.4$  events from  $Wb\bar{b}$ ,  $0.3 \pm 0.1$  events from  $Wc\bar{c}$ ,  $0.1 \pm 0.03$  events from  $Wq\bar{q}$  and  $0.1 \pm 0.03$  events from other  $W/Z$  decays. Hence we observe 2 events, compared to  $1.7 \pm 0.4$  expected from  $Wb\bar{b}$  and  $Wc\bar{c}$ , and  $0.8 \pm 0.2$  events from other backgrounds, highly dominated by top processes. We are thus consistent with the observation of  $Wb\bar{b}$ , but the hypothesis that the background by itself explains the events is not strongly disfavored.

To enhance the signal to background ratio, we tighten our tagging purity, by requiring that the jets are tagged simultaneously by the 3 tagging algorithms developed in DØ (JLIP [17], CSIP [21], Secondary vertex tagging/SVT [22]). There is correlation between the results of the 3 tagging algorithms, but the mistag background has a significantly lower probability to be tagged by the 3 tagging algorithms simultaneously than the processes with genuine  $b\bar{b}$  in the final state so its contribution becomes negligible. We observe in the simulation and in the data about 70% of common  $b$  tagging between JLIP and SVT on single tag event. We expect from the simulation that about 50% of the  $Wb\bar{b}$  events are doubly tagged simultaneously by the 3 taggers. For the  $Wc\bar{c}$  (light quarks pairs) we obtain about 30% (10%), so we also expect the contribution of  $Wc\bar{c}$  to be more reduced than the  $Wb\bar{b}$  one when requesting multiple tagging.

In figures 6,7 are shown the event displays of the 2 previous candidate events which pass also this additional requirement, i.e they have their 2 jets  $b$ -tagged by the 3 tagging algorithms. They have a clean kinematics of a  $W$  candidate recoiling against a  $b\bar{b}$  system. The transverse mass of their  $W$  candidates are 80.1 and 78.8 GeV, and their dijet invariant mass is 221 GeV and 48 GeV, respectively.

We are thus left with 2 observed events compared to  $0.6 \pm 0.2$  events expected from  $Wb\bar{b}$ , and  $0.3 \pm 0.1$  events from other standard model processes (0.1 events from  $t\bar{t}$ , 0.1 events from single-top and 0.1 events from  $Wc\bar{c}$ ). This translates into a Poissonian probability that the background explains the 2 observed events of:  $P(B) = 0.04$ , where the background  $B$  is defined as all the standard model processes except the  $Wb\bar{b}$  production, which is our signal  $S$ . Similarly, we have  $P(S + B) = 0.23$ , i.e. the hypothesis that the background alone explains the observed events is disfavored at the 2 sigma level, while the consistency that background +  $Wb\bar{b}$  explain them is preserved.

In conclusion, the only significant processes remaining to describe the 2 events which have one  $W$  candidate + 2 jets which are doubly tagged simultaneously by the 3 DØ tagging algorithms are  $Wb\bar{b}$  production and top events. The

topology, kinematics, and  $b$ -tagging characteristics of 2 of the candidates are compatible with being the first direct observation of  $Wb\bar{b}$  production. The resulting cross section is compatible with the standard model prediction.

The expected contribution from the  $b\bar{b}$  decay of a standard model Higgs boson of 115 GeV, produced in association with a  $W$ , and passing through the whole analysis chain, is shown in figure 5d, and amounts to 0.03 events. In the dijet mass window of 85–135 GeV, which has been optimized on the simulation, no events are observed. The expected standard model background (including here  $Wb\bar{b}$ ) is  $0.54 \pm 0.14$  events, and the expected  $WH$  signal is  $0.03 \pm 0.1$  event, with a signal efficiency of  $0.00154 \pm 0.00027$ . Taking these errors into account, this translates into a cross section limit for  $\sigma(p\bar{p} \rightarrow WH) \times B(H \rightarrow b\bar{b})$  of 12.4 pb at 95% C.L. limit, for a 115 GeV Higgs boson.

## IX. SUMMARY

The  $e + \cancel{E}_T + \text{jets}$  final state has been studied on  $174 \text{ pb}^{-1}$  of data taken between April 2002 and September 2003. The  $W + \geq 2$  jets channel has been analyzed in details: the ALPGEN generator, which performs at leading order in QCD and EW interactions the calculation of exact matrix elements of the elementary processes involved, describes correctly, within the current statistics, the kinematic properties (jets spectra,  $W$  transverse mass) and the production rates of these events. The shape of the dijet invariant mass,  $H_T$  and  $\Delta R$  distributions are also well described by ALPGEN.

The number of events with at least 2 jets, where at least one of the jets has been  $b$ -tagged has been measured using the Jet Lifetime Probability (JLIP) algorithm, yielding 108 events. The single  $b$ -tag production rate is consistent with the simulated expectation and the kinematic distributions of these events are well described by the simulation.

8 events have at least two jets which are  $b$ -tagged: The production rate of these double  $b$ -tagged events is in agreement with the expectation from standard model processes, within statistical and systematic errors. In this data sample, the  $Wb\bar{b}$  expectation is  $2.1 \pm 0.5$  events compared to a total of  $8.3 \pm 2.2$  events expected, in which the expectation from single-top production (0.5 events) is also taken into account. The 95% C.L. upper limit on the  $Wb\bar{b}$  cross section derived from the doubly-tagged inclusive sample, after requiring a reconstructed transverse mass for the  $W$  candidate between 25 and 125 GeV, is 20.3 pb.

Two events satisfy the additional requirement of having exactly 2  $b$ -tagged jets with  $p_T > 20$  GeV and no additional jet activity, compared to an expected  $2.5 \pm 0.7$  events, in which the  $Wb\bar{b}$  contribution is  $1.4 \pm 0.4$  events. These 2 events have their 2 jets  $b$ -tagged by the 3 tagging algorithms of DØ giving high confidence that they are  $b$  quark jets. The corresponding expectation (0.6 events from  $Wb\bar{b}$  + 0.3 events from other sources) and the kinematic properties of these candidates are compatible with being the first direct observation of  $Wb\bar{b}$  production, in agreement with the production rate expected in the standard model.

No events are found in a dijet mass window of 85–135 GeV, which is optimal for the observation of a standard model Higgs boson with mass of 115 GeV. This translates into a cross section limit for  $\sigma(p\bar{p} \rightarrow WH) \times B(H \rightarrow b\bar{b})$  of 12.4 pb at 95% C.L. limit, for a 115 GeV Higgs boson.

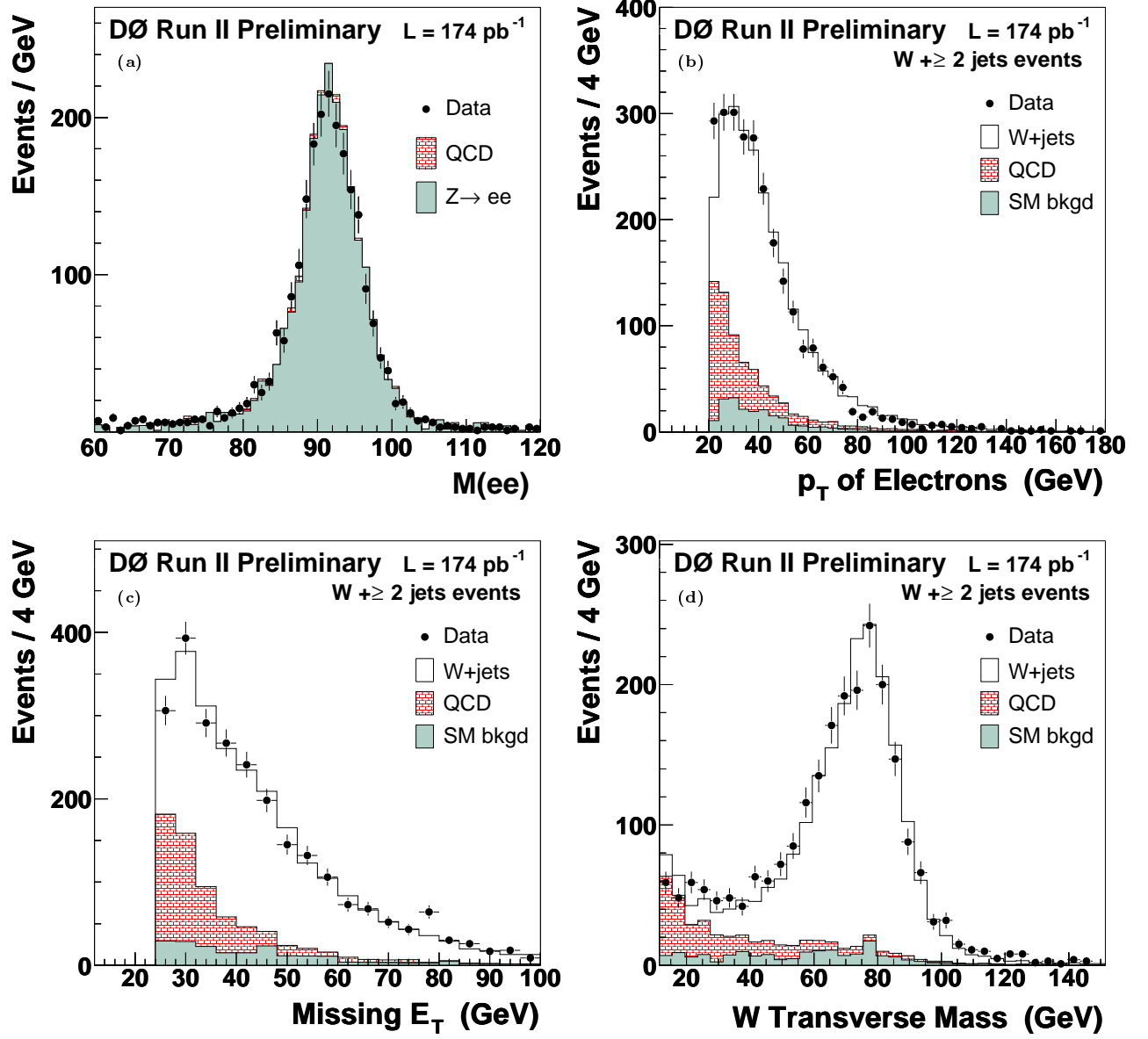
## Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à l’Energie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry for Science and Technology and Ministry for Atomic Energy (Russia), CAPES, CNPq and FAPERJ (Brazil), Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Natural Sciences and Engineering Research Council and West-Grid Project (Canada), BMBF (Germany), A.P. Sloan Foundation, Civilian Research and Development Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

- 
- [1] S. Beauceron, G. Bernardi and S. Trincas-Duvold, DØ-Note 4224
  - [2] V. Abazov, et al., in preparation for submission to Nucl. Instrum. Methods Phys. Res. A, and T. LeCompte and H.T. Diehl, “The CDF and DØ Upgrades for Run II”, Ann. Rev. Nucl. Part. Sci. **50**, 71 (2000).
  - [3] S. Abachi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994).
  - [4] Luminosity ID group: [http://www-d0.fnal.gov/phys\\_id/luminosity/data\\_access/](http://www-d0.fnal.gov/phys_id/luminosity/data_access/)



- [5] U. Bassler and G. Bernardi, Toward a coherent treatment of calorimeter energies, DØ-Note 4124
- [6] J-R. Vlimant et al., Technical description of the T42 algorithm for the calorimeter noise suppression DØ-Note 4146
- [7] G. Bernardi, E. Busato and J-R. Vlimant, Improvements from the T42 Algorithm on Calorimeter Objects Reconstruction, DØ-Note 4335
- [8] EM-ID group, Certification Results Version 5.0,  
[http://www-d0.fnal.gov/phys\\_id/emid/d0\\_private/certification/](http://www-d0.fnal.gov/phys_id/emid/d0_private/certification/)
- [9] J. Kozminski et al., Electron Likelihood in p14, DØ-Note in preparation,  
[http://www-d0.fnal.gov/jkozmins/lh\\_note\\_v1.1.pdf](http://www-d0.fnal.gov/jkozmins/lh_note_v1.1.pdf)
- [10] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, “High-energy-physics event generation with PYTHIA,” *Comput. Phys. Commun.* **135**, (2001) 238
- [11] H. L. Lai et al., Improved Parton Distributions from Global Analysis of Recent Deep Inelastic Scattering and Inclusive Jet Data, *Phys.Rev. D* **55** (1997) 1280
- [12] A. Pukhov *et al.*, COMPHEP, hep-ph/9908288 (1999)
- [13] M. Mangano et al.: ALPGEN, a generator for hard multiparton processes in hadron collisions, hep-ph/0206293
- [14] S. Beauceron et al., Systematic Studies Towards a new  $WH \rightarrow e\nu b\bar{b}$  Cross Section Limit, DØ Note 4383
- [15] Y. Fisyak and J. Womersley, DØ-Note 3191
- [16] S. Beauceron and G. Bernardi, A Search for  $Wb\bar{b}$  and  $WH$  Production at  $\sqrt{s} = 1.96$  TeV / long note, DØ Note 4394
- [17] D. Bloch et al., Jet Lifetime  $b$  Tagging, DØ Note 4069
- [18] D. Bloch et al., Performance of the JLIP  $b$ -tagger in p14, DØ Note 4348 (see also DØ Notes 4158,4159)
- [19] J. Campbell and K. Ellis, MCFM, Montecarlo for FeMtobarn processes, <http://mcfm.fnal.gov/>
- [20] I. Bertram *et al.*, Fermilab-TM-2104 (2000)
- [21] R. Demina, A. Khanov and F. Rizatdinova,  $b$  tagging with Counting Signed Impact Parameter Method, DØ-Note 4049
- [22] A. Schwartzman and M. Narain,  $b$  quark jet identification via secondary vertex reconstruction, DØ-Note 4080
- [23] A. Haas, A. Kharchilava and G. Watts, A DØ Search for Neutral Higgs Bosons at High  $\tan\beta$  in Multi-jet Events, DØ-Note 4366
- [24] The matrix method allows to derive the QCD background directly from the data, once  $p_{tight}^{loose}$  and the electron identification efficiency are known. It is described for instance in Ref. [1].



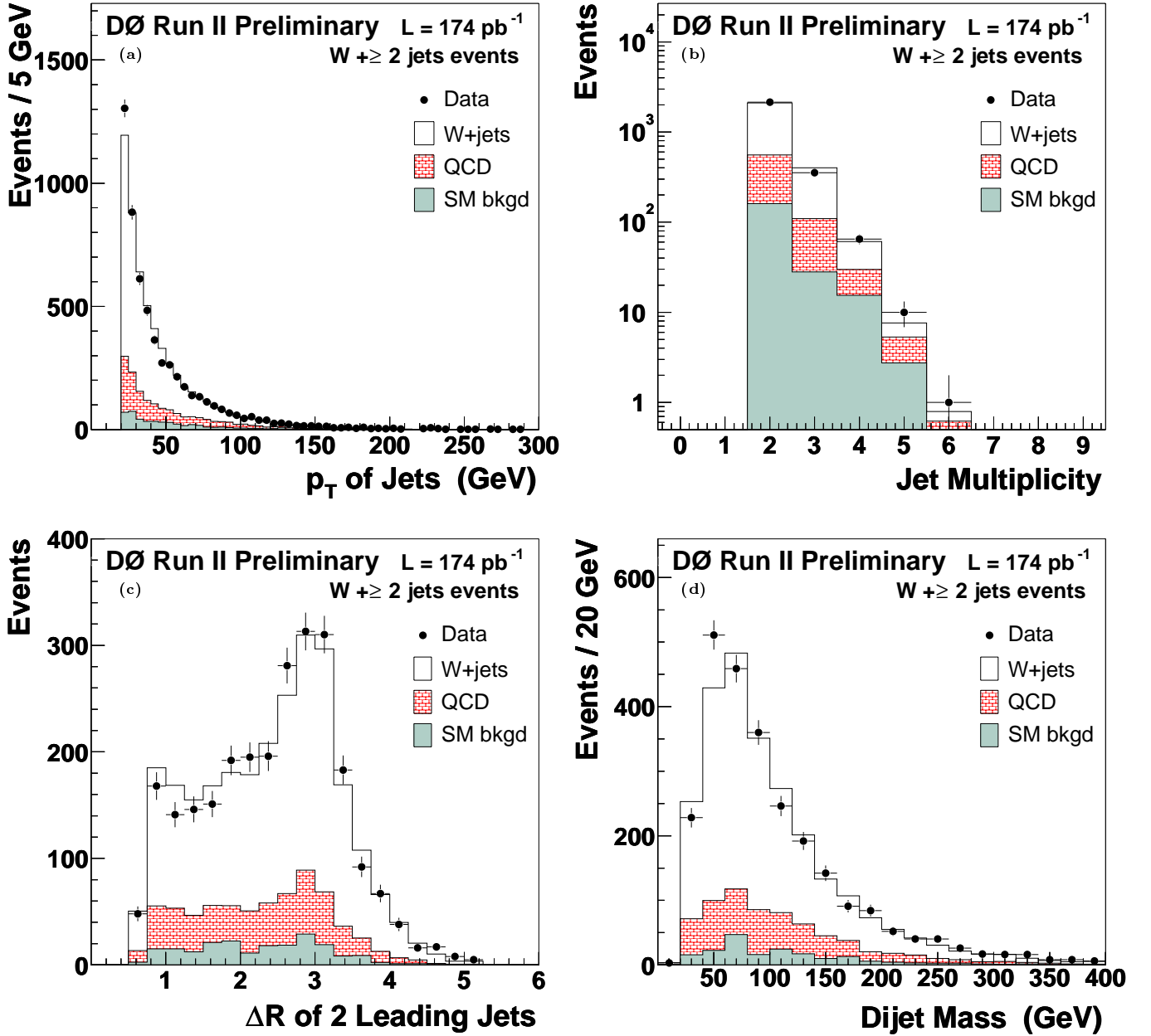


FIG. 2: Distributions obtained from the  $W$ -Dijet data sample and compared to the simulated expectation of a) transverse momentum of all the jets; b) multiplicity; c)  $\Delta R$  between the two leading jets; d) dijet invariant mass, obtained from the two leading jets. The discrepancy between the data and expectation in the  $3^{rd}$  bin is covered by the systematic errors on the data, not shown, of approximately 18%. This same remark applies to the discrepancies in a). The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the  $W$ +jets sample which is rescaled by 0.95.

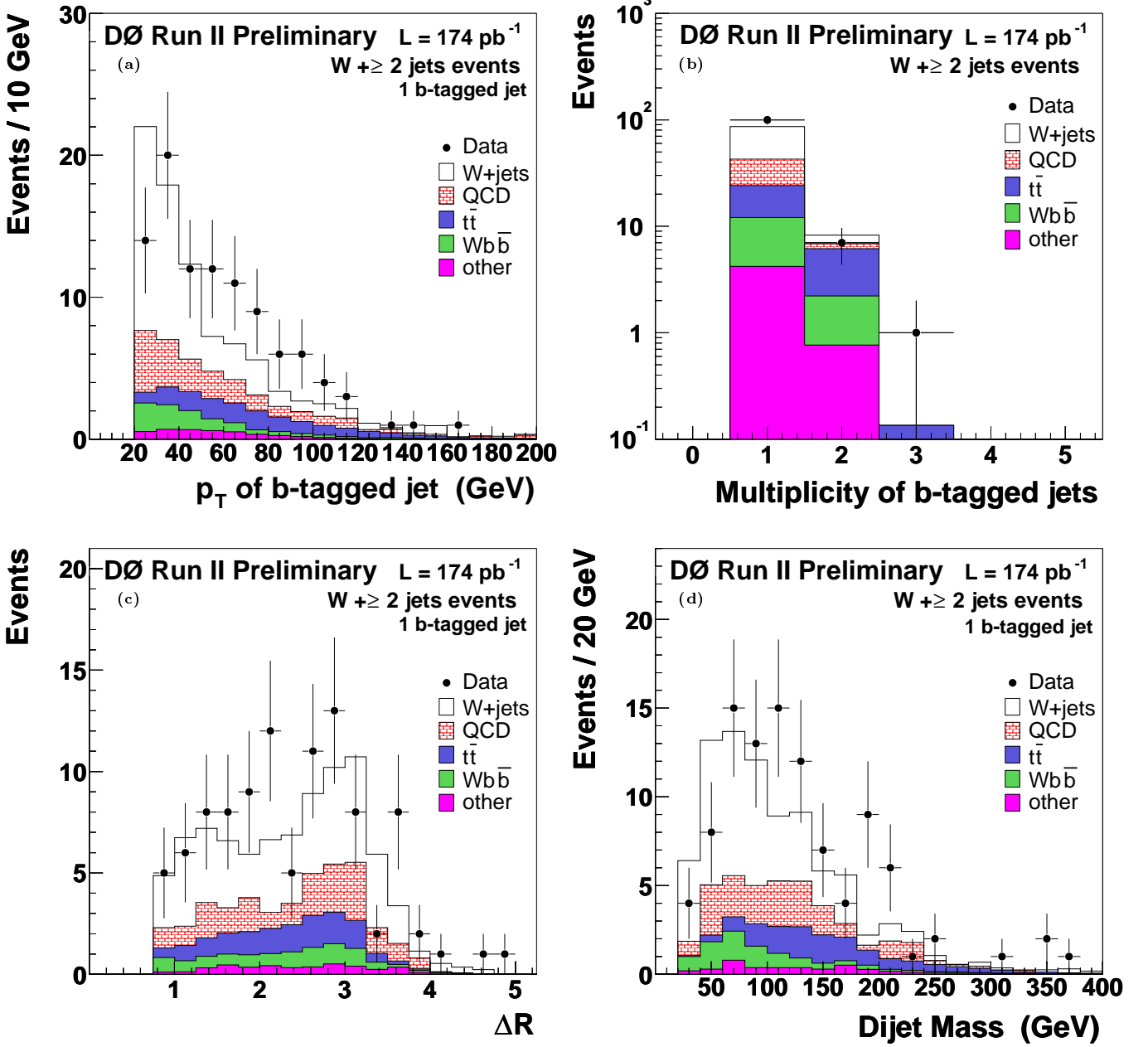


FIG. 3: a) Distributions for  $W$ -Dijet events when only one jet is  $b$ -tagged, compared to ALPGEN for the  $b$ -tagged jet transverse momentum b) the multiplicity of tagged jets; c) Distribution of  $\Delta R$  between the two leading jets when one jet is  $b$ -tagged; d) distribution of the invariant mass between the two leading jets when only one jet is  $b$ -tagged. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the  $W$ + jets sample which is rescaled by 0.95.

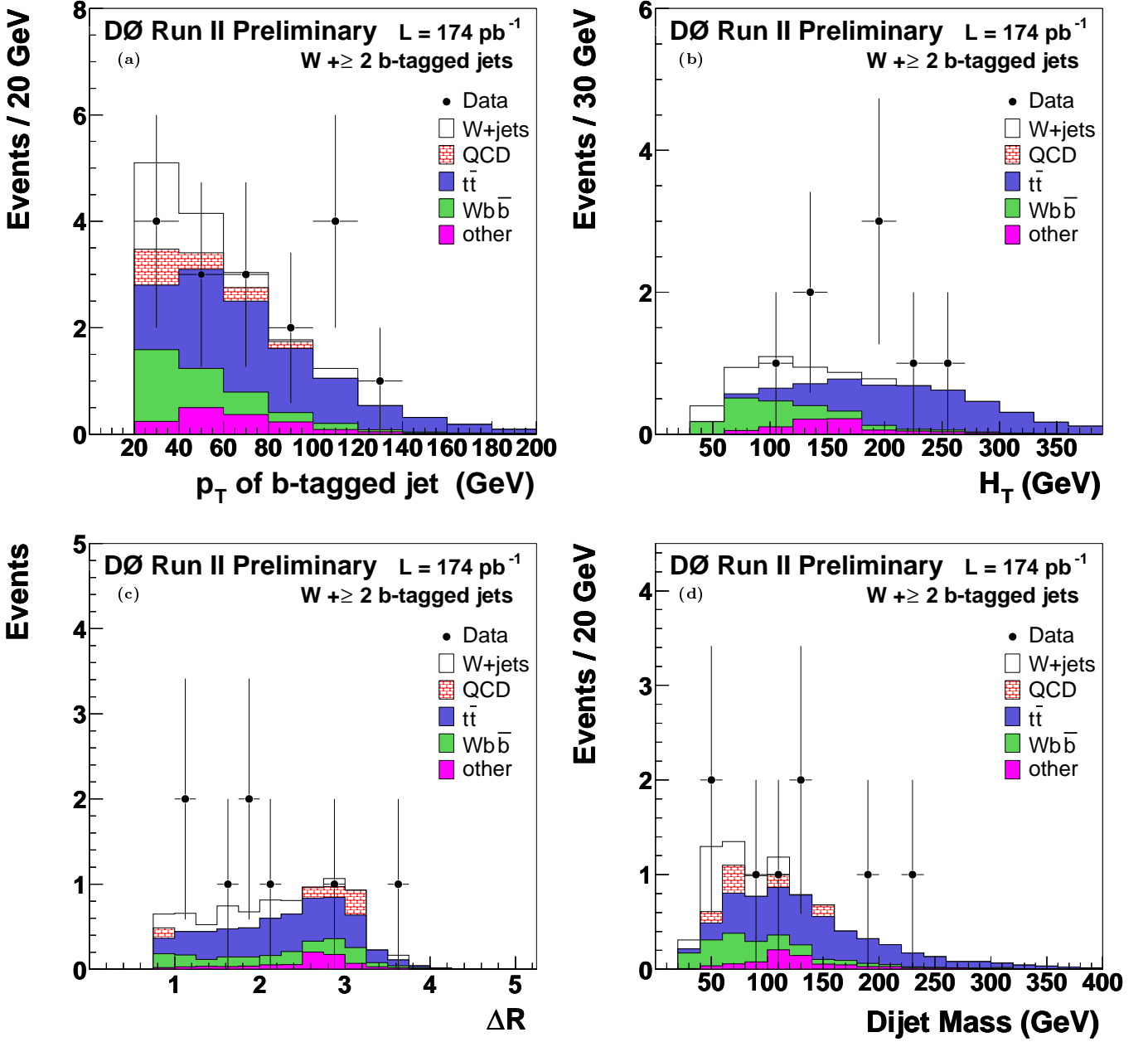


FIG. 4: Distributions for the  $W$ -Dijet events when two or more jets are  $b$ -tagged of a) the  $b$ -tagged jets momentum; b) the  $H_T$  variable; c)  $\Delta R$  between the two leading jets. d) the dijet invariant mass. The data are compared to the different simulated processes, ALPGEN being used here to simulate the  $W + n \text{ jets}$  channel. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the  $W + \text{jets}$  sample which is rescaled by 0.95.

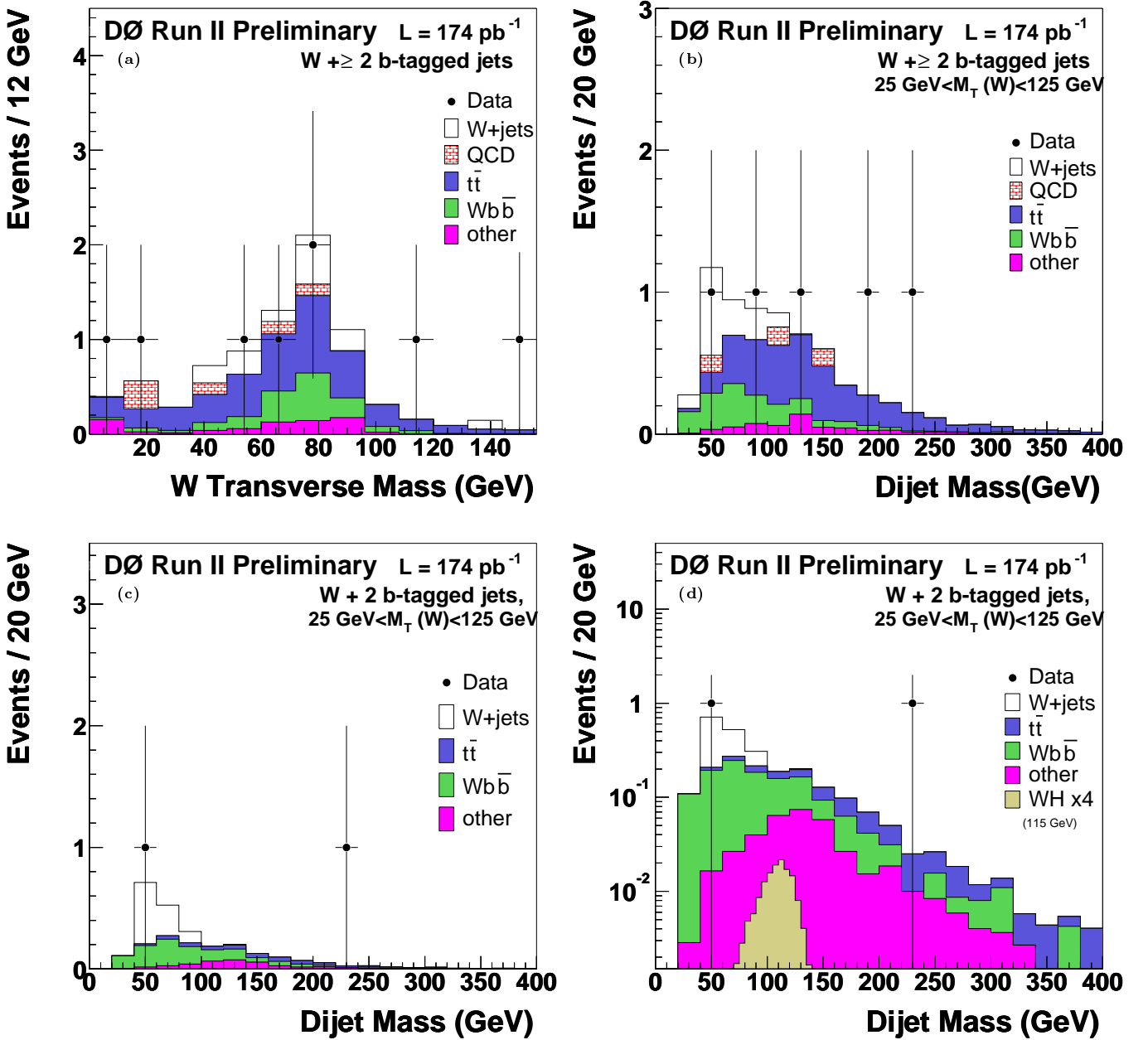


FIG. 5: Distributions for the  $W$ -Dijet events when two jets are  $b$ -tagged. The data are compared to  $Wb\bar{b}$ ,  $t\bar{t}$ ,  $W$ +jets and other smaller expectations. Note that  $W$ +jets also contains  $Wb\bar{b}$  produced by parton shower ( $0.4 \pm 0.2$  events in figures c,d for instance, see text for details of the  $Wb\bar{b}$  fraction in  $W$ +jets for the other figures). The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the  $W$ + jets sample which is rescaled by 0.95. The “other” backgrounds are dominated by single-top production. a)  $W$  transverse mass; b) dijet invariant mass of the events having  $25 < m_T < 125$  GeV; c) same as b), but requiring that the events have exactly 2 jets; d) same as c), but in logarithmic scale; also shown is the contribution expected from the 2  $b$  jets originating from the decay of a 115 GeV Higgs Boson produced via  $p\bar{p} \rightarrow WH$ , with a finer binning (5 GeV) and multiplied by 4. The mean / RMS of this expected distribution fitted by a Gaussian in the optimal mass window 85–135 GeV is 108.5 GeV / 15.1 GeV, i.e. a relative resolution of 14%. The uncertainty on this resolution has been estimated to be 12% [23] but has no influence on the quoted cross section limit, since no data events are observed in a window twice as large as the optimal mass window, which is expected to contain 98.6% of the Higgs events.

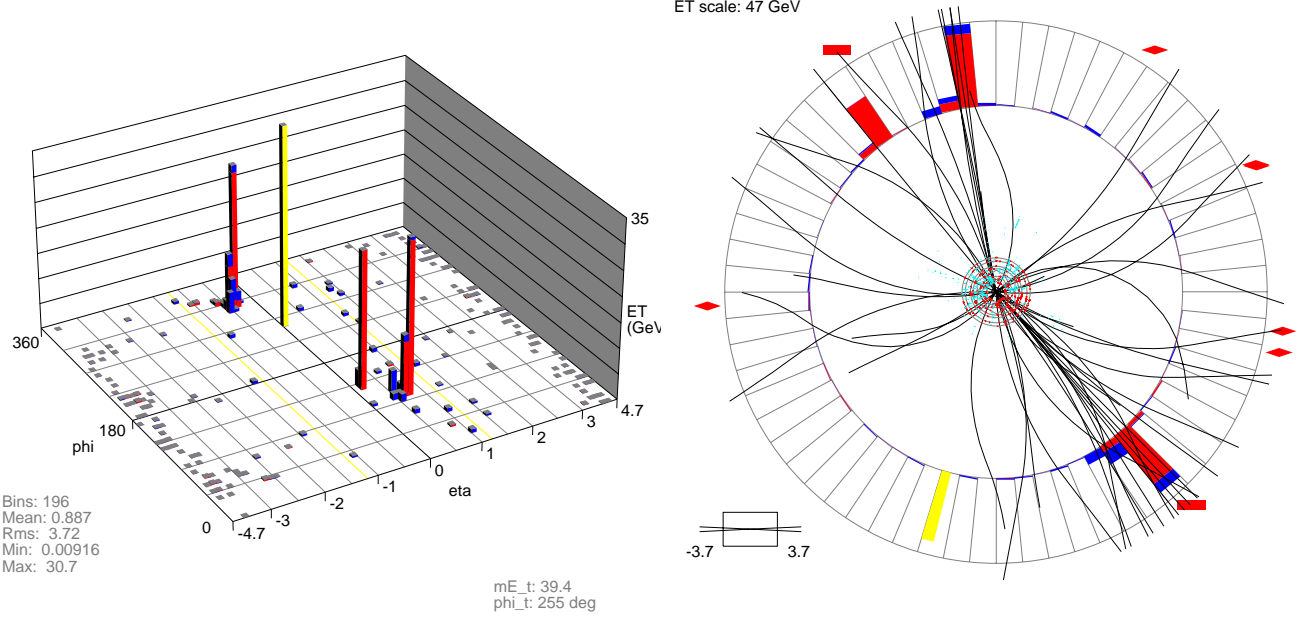


FIG. 6: Event Display of one of the 2 events in the  $W$  mass window having exactly 2 jets with  $p_T > 20$  GeV, both  $b$ -tagged by the 3 DØ tagging algorithms and displaying no additional jet activity; a) Lego plot: the highest “tower” represent  $\cancel{E}_T$ , the electron is the “tower” at  $\eta \simeq 0.2$  and  $\varphi \simeq 145^\circ$ . The 2 jets are the 2 remaining sets of towers. b) Transverse view:  $\cancel{E}_T$  is represented by the single yellow tower pointing to the bottom, the electron points top left.

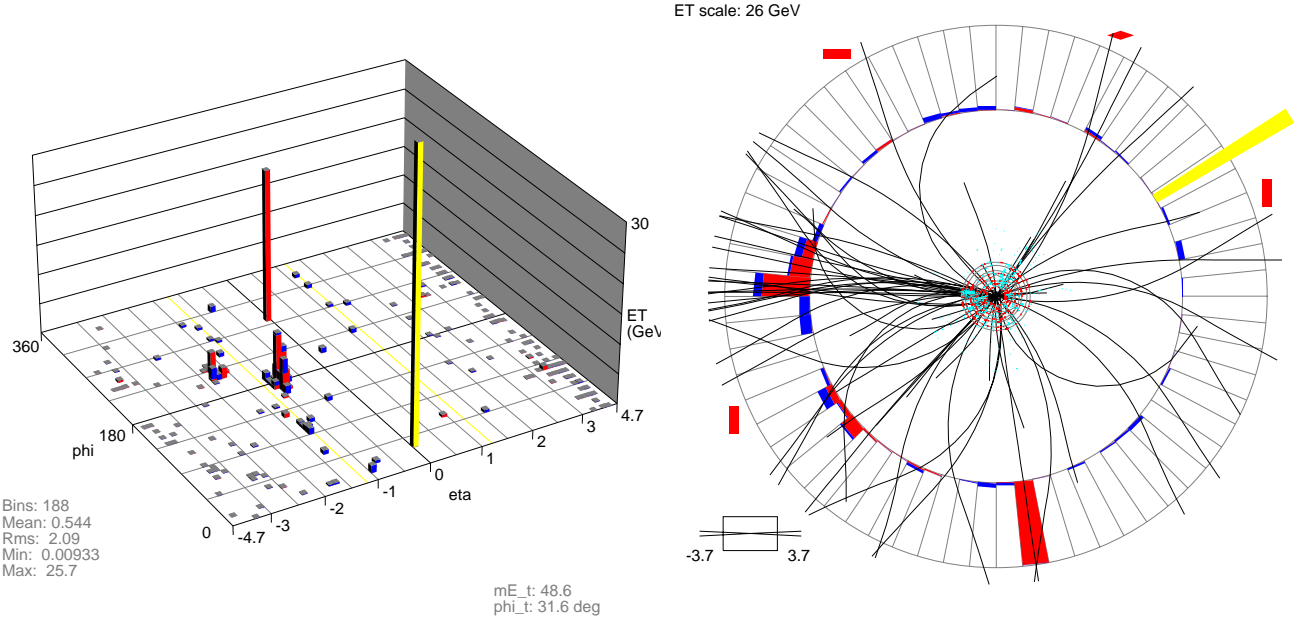


FIG. 7: Event Display of the other event in the  $W$  mass window having exactly 2 jets with  $p_T > 20$  GeV, both  $b$ -tagged by the 3 DØ tagging algorithms and displaying no additional jet activity; a) Lego plot: the highest “tower” represent  $\cancel{E}_T$ , the electron is the second highest “tower”. The 2 jets are the 2 remaining sets of towers. b) Transverse view:  $\cancel{E}_T$  is represented by the single yellow tower pointing to the top right, the electron points to the bottom.